D104–D108 *Nucleic Acids Research, 2004, Vol. 32, Database issue*
DOI: 10.1093/nar/gkh102

The tmRNA website: reductive evolution of tmRNA in plastids and other endosymbionts

Pulcherie Gueneau de Novoa and Kelly P. Williams[1,*]

Centro de Biofisica y Bioquimica, Instituto Venezolano de Investigaciones Cientificas, Altos de Pipe, MI, Venezuela and [1]Department of Biology, Indiana University, Bloomington, IN 47405, USA

Received September 23, 2003; Revised and Accepted October 13, 2003 DDBJ/EMBL/GenBank accession nos+

ABSTRACT

tmRNA combines tRNA- and mRNA-like properties and ameliorates problems arising from stalled ribosomes. Research on the mechanism, structure and biology of tmRNA is served by the tmRNA website (http://www.indiana.edu/~tmrna), a collection of sequences, alignments, secondary structures and other information. Because many of these sequences are not in GenBank, a BLAST server has been added; another new feature is an abbreviated alignment for the tRNA-like domain only. Many tmRNA sequences from plastids have been added, five found in public sequence data and another 10 generated by direct sequencing; detection in early-branching members of the green plastid lineage brings coverage to all three primary plastid lineages. The new sequences include the shortest known tmRNA sequence. While bacterial tmRNAs usually have a lone pseudoknot upstream of the mRNA segment and a string of three or four pseudoknots downstream, plastid tmRNAs collectively show loss of pseudoknots at both postions. The pseudoknot-string region is also too short to contain the usual pseudoknot number in another new entry, the tmRNA sequence from a bacterial endosymbiont of insect cells, *Tremblaya princeps*. Pseudoknots may optimize tmRNA function in free-living bacteria, yet become dispensible when the endosymbiotic lifestyle relaxes selective pressure for fast growth.

INTRODUCTION

tmRNA helps solve problems associated with stalled ribosomes and is essential in some but not all bacteria (1). It contains a tRNA-like domain that is charged with alanine but has no anticodon; instead, the corresponding stem (P2) is extended and capped by a large looping domain of RNA (2,3). Within this loop is a reading frame that is translated in an unusual way. tmRNA engages ribosomes that have stalled, e.g. at the end of an mRNA with no in-frame stop codon; the alanine moiety attached to tmRNA is added to the nascent protein, and the ribosome is directed to a particular triplet on tmRNA, termed the resume codon, from which the tmRNA reading frame is translated. Consequently, the ribosome is freed, and the nascent protein is tagged with a peptide sequence that signals its degradation.

The P2 stem is long and probably further stabilized by coaxial stacking with an abutting pseudoknot. If P2 persists throughout translation, tmRNA presents its reading frame to the ribosome in the unusual form of a looped mRNA, which could cause topological problems for the translation of tmRNA. The problem may be ameliorated by pseudoknots in the loop, which could open to relieve strain during translation, then reclose to take up slack. In the loop RNA, there is usually a lone pseudoknot abutting P2 and a string of three pseudoknots (four in cyanobacteria) at the 3′ end (Fig. 1). A second solution to the potential topological problems posed by a looped mRNA is simply to break open the loop. This actually occurs in the two-piece tmRNAs that are produced from permuted genes in two bacterial lineages, the α-proteobacteria and a group of cyanobacteria (4,5). In both these lineages, the looped domain is opened, and correspondingly, the number of pseudoknots in the string dwindles, from three or four, to one.

Despite their conservation in most one-piece tmRNAs, none of the pseudoknots of the string were absolutely required for *in vitro* translation of *Escherichia coli* tmRNA (6). Here we describe natural cases where pseudoknot number in the string is reduced in one-piece tmRNAs. This occurs in certain residents of eukaryotic cells: the photosynthetic plastids and the endosymbiotic bacterium *Candidatus Tremblaya princeps*. Difficulty in finding these pseudoknots had been noted when tmRNA sequences were first identified in plastids (7), but several tmRNA genes newly found or sequenced here make the case more strongly, especially that from *Cyanidium caldarium*, in which the segment is far shorter than usual.

While deletion of pseudoknots in the string was tolerated in *E.coli* tmRNA, disruption or deletion of the lone pseudoknot upstream of the reading frame was deleterious (6,8). In contrast, even this pseudoknot can be lost in plastid tmRNA, as clearly evidenced by the sequence from *Cyanidioschyzon merolae* (Fig. 1).

*To whom correspondence should be addressed at 1001 E. Third Street, Bloomington, IN 47405, USA. Tel: +1 812 856 5697; Fax: +1 812 855 6705; Email: kellwill@indiana.edu

+AY313266–AY313270, AF169625, AF169626 and AF550350–AF550357

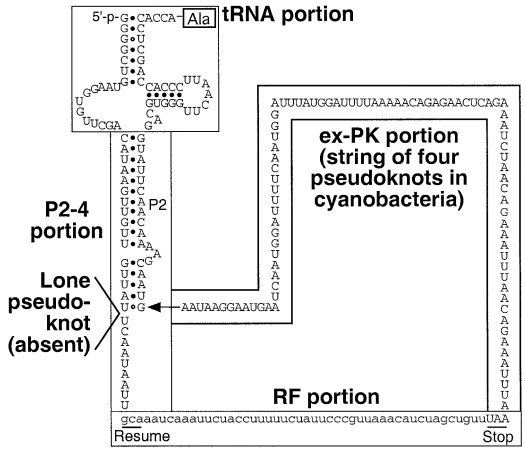

Figure 1. The shortest tmRNA, from the *C.merolae* plastid, showing the four portions analyzed in Table 1. The complete loss of the lone upstream pseudoknot is idiosyncratic.

It may be that while pseudoknots improve one-piece tmRNA function, as evidenced by their conservation throughout free-living bacteria, they are not essential and can be lost when the endosymbiotic lifestyle relaxes selective pressure for optimal tmRNA function.

THE tmRNA WEBSITE

The tmRNA website was established in 1997 to facilitate research on tmRNA biology, structure, evolution and mechanism of action. It presents tmRNA sequences and alignments, color-coded to show secondary structure, together with careful documentation and other information. Public databases are vigorously monitored for new tmRNA sequences, and updates are frequent. Records of variant sequences from different isolates of the same species are maintained. As of October 2003, the website contained 366 non-duplicate sequences from 306 species. A complete list of all sequences in the database, in FASTA format, is available for download. A sequence alignment for SmpB, the protein cofactor of tmRNA, is also provided.

Since many of the tmRNA sequences are not available at GenBank, coming from diverse incomplete genome projects, a BLAST server has been added. Another new feature is a truncated alignment, of only the tRNA-like domain and the long pairing P2 that protrudes from it. New tmRNA sequences from plastids and other endosymbionts exhibit reductive evolution.

PROGRESS AMONG THE PLASTIDS

Most lines of evidence point to the origin of all plastids from a single endosymbiotic cyanobacterium, although there is no stand-out among extant cyanobacteria as a representative of this progenitor (9–11). Previous identification of the tmRNA gene (*ssrA*) in plastids (7) and the new identifications described below are based on several key features: the distinctive tRNA-like domain from which the long stem P2 emerges, the tag reading frame encoding the characteristic hydrophobic C-terminal residues, and usually the lone pseudoknot upstream of the reading frame (Supplementary Material). The predicted tag sequences show strong relationships to those of the cyanobacteria (Table 1), supporting the expected vertical transmission of *ssrA* from a cyanobacterial ancestor. This conservation suggests that the genes produce active tmRNAs, and *ssrA* from *Cyanophora paradoxa* has been shown to produce an RNA with the predicted 5′ end (15). However tmRNA activity has not been directly demonstrated in any plastid.

There are three primary plastid lineages (green, red, and glaucocystophyte), although plastid phylogeny is complicated by events of secondary (and tertiary) endosymbiosis, in which eukaryotes have engulfed other, plastid-containing eukaryotes (9,16). tmRNA sequences have beenidentified previously in the latter two primary lineages (7). We have now identified the first known tmRNA sequences from the green plastid lineage by homology- and pattern-searching in the plastid genomes from *Mesostigma viride* and *Nephroselmis olivacea*. The *N.olivacea* plastid has the only genome yet known to contain two copies of *ssrA*: these occur in two large inverted repeat segments. While this paper was in preparation, the *M.viride* sequence was identified as tmRNA by another group, without comment on the phylogenetic significance (17). *M.viride* has been placed alternatively in a green algal lineage emerging before the divergence of the two main green plant lineages (streptophytes and chlorophytes) or as a very early-diverging group within the streptophytes (18,19). *N.olivacea* is an early-diverging chlorophyte. We have not identified *ssrA* in the complete genome of any other green plastid, and if indeed absent, it was lost at least twice, independently in the chlorophytes and streptophytes.

Most plastid tmRNA sequences are known from the red plastid lineage. We identified *ssrA* in three additional red plastid genomes by homology- and pattern-searching, and exploited the conservation at both its termini (2) to amplify and sequence it from 10 others (Table 1). A novel feature is proposed for the three available sequences of the Bangiales/Florideophycidae group of red algae, a stem–loop inserted in the 3′ portion of the long pairing P2 (Supplementary Material). An appealing feature of this proposal is that it allows P2 and P3 to abut, probably with coaxial stacking, as for most other plastid and bacterial tmRNAs. A similar structure can be proposed for the brown alga *Pylaiella littoralis*, awaiting confirmation with sequences from other brown algae. Idiosyncratic branches have been identified at other sites in tmRNAs, in P6 of β-proteobacteria and in P5 of cyanobacteria (2,20,21), but not previously in P2.

PSEUDOKNOT LOSS IN ENDOSYMBIONT tmRNAs

tmRNA sequences of plastids and cyanobacteria were split into four portions for analysis (Fig. 1; Table 1). The most irregular portion was the segment (ex-PK) downstream of the reading frame, which in bacteria contains the string of pseudoknots. Its length was extremely variable, even within subgroups of related plastids [in contrast, the variability of the reading frame (RF) portion length was ascribed to a single insertion event for the diatoms], with a mean only two-thirds of that for cyanobacteria. G+C content was distinctly lower for the ex-PK segment than for the RF or P2–4 portions in plastids, but was similar among these portions in cyanobacteria.

Table 1. Phylogeny of plastids with tmRNA sequence available, and statistics for portions of tmRNA

tmRNA portion[a]	Length (nt) tRNA	P2–4	RF	exPK	G+C content (%) tRNA	P2–4	RF	exPK	gnm	Predicted encoded tag
GREEN PLASTIDS										
Mesostigma viride[b]	47	112	48	149	57.4	32.1	33.3	22.8	30.2	ANNILPFNRK----------TAVAV
Nephroselmis olivacea[b]	47	117	48	112	46.8	47.0	50.0	40.2	42.1	ANQILPFSRR----------VAVAA
GLAUCOCYSTOPHYTE PLASTIDS										
Cyanophora paradoxa	47	111	48	85	55.3	27.9	35.4	15.3	30.5	ATNIVRFNRK----------AAFAV
RED PLASTIDS										
Cyanidiales red algae										
Cyanidium caldarium[b]	48	106	54	43	54.2	34.0	29.6	18.6	32.7	ANNIIEISNIRK--------PALVV
Cyanidioschyzon merolae[b]	47	44	51	90	57.4	25.0	31.4	24.4	37.6	ANQILPFSIPVK---------HLAV
Chromists (2° endosymbiosis)										
Cryptophytes										
Guillardia theta	47	112	51	117	53.2	26.8	29.4	23.9	33.0	ASNIVSFSSKR---------LVSFA
Rhodomonas salina[c,e]	ND	110	45	147	ND	31.8	31.1	25.2	ND	ANNIVPFSRK----------VALV
Heterokonts										
diatoms										
Amphora coffaeformis[c,e]	ND	110	72	99	ND	28.2	22.2	19.2	ND	ATIITWFISKIINRNACS--LQFVV
Thalassiosira weissflogii	47	108	78	112	48.9	24.1	24.4	23.2	ND	ANNIIPFIFKAVKTKKEAMALNFAV
Thalassiosira pseudonana[b]	47	107	78	109	51.1	20.6	26.9	18.3	ND	ANNIMPFMFNVVKTNRSLTTLNFAV
Odontella sinensis	47	111	78	132	59.6	24.3	20.5	24.2	31.8	ANNLISSVFKSLSTKQNSLNLSFAV
Skeletonema costatum[c,e]	ND	109	78	109	ND	25.7	23.1	15.6	ND	ANNIMSFIFKTVTPKNHLNVLSFAV
Fragillaria pinnata[c,e]	ND	109	75	101	ND	28.4	21.3	26.7	ND	ANNIIPFHFKTVNFNNSNL-LQFAA
bolidophytes										
Bolidomonas pacifica[d,e]	ND	108	48	142	ND	17.6	27.1	19.7	ND	ANNILAFNRK----------SLSFA
brown algae										
Pylaiella littoralis[d,e]	ND	103	45	200	ND	17.5	26.7	19.5	ND	ANNIMSFNK-----------NQVFA
Haptophytes										
Pavlova lutheri	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Prymnesium parvum[c,e]	ND	102	45	92	ND	25.5	22.2	25.0	ND	ANNILSFNT-----------KLALA
Non-cyanidiales red algae										
Stylonema/Bangiopsis										
Rhodosorus marinus[c,e]	ND	ND	54	109	ND	ND	33.3	25.7	ND	ANNILKFFTKSP--------VVAFA
Bangiales/Florideophycidae										
Porphyra purpurea	46	121	48	105	58.7	22.3	27.1	24.8	33.0	AENNIIAFSR----------KLAVA
Gigartina papillata[d,e]	ND	125	48	112	ND	28.8	22.9	16.1	ND	AKHQIVPFSK----------RIIVV
Prionitis lanceolata[d,e]	ND	133	48	131	ND	28.6	27.1	22.9	ND	AKHQILPLSR----------KIALA
Mean for plastids	47	108	57	115	54.3	27.2	28.3	22.6	33.9	
SD (% of mean)	1.0	16.0	23.5	27.3	7.9	24.1	23.8	24.3	11.9	
Mean for 12 cyanobacteria[f]	47	115	49	177	58.9	47.2	45.8	48.0	43.7	ANNIVPFARKQ---------VAALA
SD (% of mean)	0.0	0.9	4.8	3.0	6.5	4.7	11.6	9.2	7.4	

Red plastid groupings are based on (12,13). ND, not determined, e.g. tRNA portion sequence is incomplete when generated by PCR.

[a]Four tmRNA portions: tRNA, the tRNA-like domain excluding the usually uncoded 3′ CCA tail; P2–4, the segments containing long stem P2, the lone upsteam pseudoknot and extending to the resume codon; RF, the reading frame segment, from the resume codon to the stop codon inclusive; ex-PK, the downstream segment between the stop codon and P2 (Fig. 1); gnm, genome.

[b]Identified using BLAST (14) or PatScan (R. Overbeek) searches.

[c]Total DNA was prepared from cells harvested from axenic cultures purchased from the Center for Culture of Marine Phytoplankton (West Boothbay Harbor, ME), using cetyltrimethylammonium bromide as in (2).

[d]Genomic DNA samples were kind gifts from L. Guillou (University of Copenhagen) (*B.pacifica*), S. Loiseaux-DeGoër (Station Biologique de Roscoff, France) (*P.littoralis*) and P. Keeling (University of British Columbia) (*G.papillata* and *P.lanceolata*).

[e]*ssrA* was amplified by PCR (2) and sequenced in both directions using the same primers (GenBank accession numbers AF169625–AF169626 and AF550350–AF550357).

[f]One-piece tmRNAs from the following cyanobacteria were evaluated: *Nostoc* sp. PCC 7120 and *Nostoc punctiforme*, *Fremyella diplosiphon*, *Plectonema boryanum*, *Trichodesmium erythraeum*, *Oscillatoria* spp. PCC 6304 and PCC 7515, *Chroococcidiopsis* sp. PCC 6712, *Synechocystis* sp. PCC 6803, *Thermosynechococcus elongatus*, *Synechococcus* spp. PCC 7002 and PCC 6301. The *T.erythraeum* tag sequence is shown.

A priori, the shorter mean length and extreme nucleotide bias toward A and U in the plastid ex-PK segment would tend to disfavor the full four-pseudoknot secondary structure found in cyanobacteria. Plastid ex-PK segment sequences are difficult to align, with partial success only for (i) the green algae, (ii) the diatoms and (iii) the Bangiales/Florideophycidae red algae (Supplementary Material). Base-pair covariation in plastid ex-PK segments currently supports one pseudoknot only, at the 3′ end in the green algae. Some plastids may have lost all pseudoknots of the ex-PK segment: *Cyanidium caldarium* makes a compelling case. Not counting the last 3 nt, which are unpaired in bacteria, its ex-PK segment contains only 40 nt, 22 of which are U, and only eight of which are G or C. Allowing stems with as few as three contiguous Watson–Crick or G:U pairs, the possible pseudoknots in this segment can be enumerated at 61; 11 have 7 or 8 bp total in the two stems, but the rest have the minimum 6 bp, and one of the latter is the only one with any G:C base pairs. *C.caldarium* is a thermophile, living at temperatures of over 45°C (22), making pseudoknots with very short A:U-rich stems even less likely.

The tmRNA encoded by the *C.merolae* plastid would be the shortest known, at 235 nt (Fig. 1). Although its ex-PK segment is quite short, what is remarkable is its complete loss of the lone pseudoknot upstream of the reading frame; the corresponding pseudoknot in *E.coli* is thought to be a strong determinant of tmRNA function (8).

Another endosymbiont tmRNA sequence with reduced pseudoknot number, from a bacterium, was identified in a BLAST search at GenBank. *T.princeps* is a β-proteobacterial endosymbiont living within specialized cells of mealybugs, that is notable for its ability to harbor other bacteria within its cytoplasm (23,24). We amplified and sequenced (GenBank accession numbers AY313266–AY313270) *ssrA* from DNA samples of *T.princeps* from five additional mealybug hosts, kindly provided by P. Baumann (University of California, Davis), so that all six of the main *T.princeps* subgroups (25) were represented. A regular tRNA-like domain and P2 can be detected (Supplementary Material). One reading frame in the sequence is most likely to be the tag reading frame because it ends with the characteristic codons; however, the resume codon is uncertain. The most striking feature of the *T.princeps* sequence is its short ex-PK segment (54–55 nt), which is >150 nt for all other known bacteria. For some *T.princeps* strains a pseudoknot can be drawn in this segment, but it is not conserved in all strains. In any case the segment is clearly too short to contain the usual number of pseudoknots.

REDUCTIVE EVOLUTION OF ENDOSYMBIONT tmRNAs

We have shown that two lineages of endosymbionts have independently lost tmRNA pseudoknots, although we note that many endosymbiotic bacteria retain normal pseudoknot numbers. These losses could be considered part of the typical reductive evolution of endosymbiont genomes (26). Some of the four pseudoknots in the ex-PK segment of the presumable cyanobacterial ancestor may have been lost before the divergence of the three plastid lineages, during the genomic upheavals of early plastid domestication. This segment of *ssrA* would then be vulnerable to mutations accruing with varied histories among different plastid lineages. Even more drastic change was noted in the mitochondrial homolog of tmRNA, which is little more than a tRNA-like domain with a P2 stem; however, this case is not particularly comparable to the plastids and *Tremblaya* because the mitochondrial gene (i) is permuted, and (ii) lacks a tag reading frame, and has therefore lost full tmRNA function (4).

The pseudoknot string may solve the topological difficulty of translating a relatively small, looped mRNA domain: pseudoknots could open and reclose to dynamically regulate the size of the reading frame loop during translation. Pseudoknot number is reduced in both cases where the mature tmRNA takes a two-piece form and the mRNA domain loop is opened (4,5). Among one-piece tmRNAs, the conservation of pseudoknots throughout free-living bacteria attests to their contribution to tmRNA function, yet their occasional loss in endosymbionts suggests that such contributions are dispensible, in agreement with an *in vitro* study (6). With the relaxed selective pressure for fast growth that plastids and *Tremblaya* enjoy as endosymbionts, suboptimal tmRNA activity may become tolerable. Another possibility is that pseudoknots serve mainly to protect tmRNA from ribonuclease activity, which may be reduced in these endosymbionts.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank P. Baumann, L. Guillou, S. Loiseaux-DeGoër and P. Keeling for DNA samples. This work was supported by an NIH grant to K.P.W.

REFERENCES

1. Keiler,K.C., Waller,P.R. and Sauer,R.T. (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, **271**, 990–993.
2. Williams,K.P. and Bartel,D.P. (1996) Phylogenetic analysis of tmRNA secondary structure. *RNA*, **2**, 1306–1310.
3. Felden,B., Himeno,H., Muto,A., McCutcheon,J.P., Atkins,J.F. and Gesteland,R.F. (1997) Probing the structure of the *Escherichia coli* 10Sa RNA (tmRNA). *RNA*, **3**, 89–103.
4. Keiler,K.C., Shapiro,L. and Williams,K.P. (2000) tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: A two-piece tmRNA functions in *Caulobacter*. *Proc. Natl Acad. Sci. USA*, **97**, 7778–7783.
5. Gaudin,C., Zhou,X., Williams,K.P. and Felden,B. (2002) Two-piece tmRNA in cyanobacteria and its structural analysis. *Nucleic Acids Res.*, **30**, 2018–2024.
6. Nameki,N., Tadaki,T., Himeno,H. and Muto,A. (2000) Three of four pseudoknots in tmRNA are interchangeable and are substitutable with single-stranded RNAs. *FEBS Lett.*, **470**, 345–349.
7. Williams,K.P. and Bartel,D.P. (1998) The tmRNA Website. *Nucleic Acids Res.*, **26**, 163–165.
8. Nameki,N., Felden,B., Atkins,J.F., Gesteland,R.F., Himeno,H. and Muto,A. (1999) Functional and structural analysis of a pseudoknot upstream of the tag-encoded sequence in *E. coli* tmRNA. *J. Mol. Biol.*, **286**, 733–744.
9. Bhattacharya,D. and Medlin,L. (1995) The phylogeny of plastids: a review based on comparisons of small-subunit ribosomal RNA coding regions. *J. Phycol.*, **31**, 489–498.
10. Palmer,J.D. and Delwiche,C.F. (1998) The origin and evolution of plastids and their genomes. In Soltis,D.E., Soltis,P.S. and Doyle,J.J. (eds), *Molecular Systematics of Plants II*. Kluwer Academic Publishers, Boston, pp. 375–409.
11. Turner,S., Pryer,K.M., Miao,V.P. and Palmer,J.D. (1999) Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J. Eukaryot. Microbiol.*, **46**, 327–338.
12. Yoon,H.S., Hackett,J.D., Pinto,G. and Bhattacharya,D. (2002) The single, ancient origin of chromist plastids. *Proc. Natl Acad. Sci. USA*, **99**, 15507–15512.
13. Daugbjerg,N. and Guillou,L. (2001) Phylogenetic analyses of Bolidophyceae (Heterokontophyta) using *rbcL* gene sequences support their sister group relationship to diatoms. *Phycologia*, **40**, 153–161.
14. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
15. Gimple,O. and Schön,A. (2001) *In vitro* and *in vivo* processing of cyanelle tmRNA by RNase P. *Biol. Chem.*, **382**, 1421–1429.
16. Cavalier-Smith,T. (1981) Eukaryote kingdoms: seven or nine? *BioSystems*, **14**, 461–481.
17. Zwieb,C., Gorodkin,J., Knudsen,B., Burks,J. and Wower,J. (2003) tmRDB (tmRNA database). *Nucleic Acids Res.*, **31**, 446–447.
18. Karol,K.G., McCourt,R.M., Cimino,M.T. and Delwiche,C.F. (2001) The closest living relatives of land plants. *Science*, **294**, 2351–2353.
19. Turmel,M., Otis,C. and Lemieux,C. (2002) The complete mitochondrial DNA sequence of *Mesostigma viride* identifies this green alga as the earliest green plant divergence and predicts a highly compact mitochondrial genome in the ancestor of all green plants. *Mol. Biol. Evol.*, **19**, 24–38.

20. Felden,B., Massire,C., Westhof,E., Atkins,J.F. and Gesteland,R.F. (2001) Phylogenetic analysis of tmRNA genes within a bacterial subgroup reveals a specific structural signature. *Nucleic Acids Res.*, **29**, 1602–1607.
21. Williams,K.P. (2002) Descent of a split RNA. *Nucleic Acids Res.*, **30**, 2025–2030.
22. Doemel,W.N. and Brock,T.D. (1970) The upper temperature limit of *Cyanidium caldarium*. *Arch. Mikrobiol.*, **72**, 326–332.
23. von Dohlen,C.D., Kohler,S., Alsop,S.T. and McManus,W.R. (2001) Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature*, **412**, 433–436.
24. Baumann,L., Thao,M.L., Hess,J.M., Johnson,M.W. and Baumann,P. (2002) The genetic properties of the primary endosymbionts of mealybugs differ from those of other endosymbionts of plant sap-sucking insects. *Appl. Environ. Microbiol.*, **68**, 3198–3205.
25. Thao,M.L., Gullan,P.J. and Baumann,P. (2002) Secondary (γ-Proteobacteria) endosymbionts infect the primary (β-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts. *Appl. Environ. Microbiol.*, **68**, 3190–3197.
26. Andersson,S.G. and Kurland,C.G. (1998) Reductive evolution of resident genomes. *Trends Microbiol.*, **6**, 263–268.